

# Report from the Metadata Working Group



Giuseppe Andronico (Italy)  
Paul Coddington (Australia)  
Robert Edwards, Jim Simone (USA)  
Balint Joo, Chris Maynard (UK)  
Dirk Pleiter (Germany)  
Tomoteru Yoshie (Japan, Convener)

# Introduction

Sharing gauge configurations requires

- ❑ Standardised description of the configuration (=metadata)
  - XML documents which conform to a XML schema
- ❑ Standards on binary file format

**Status:**

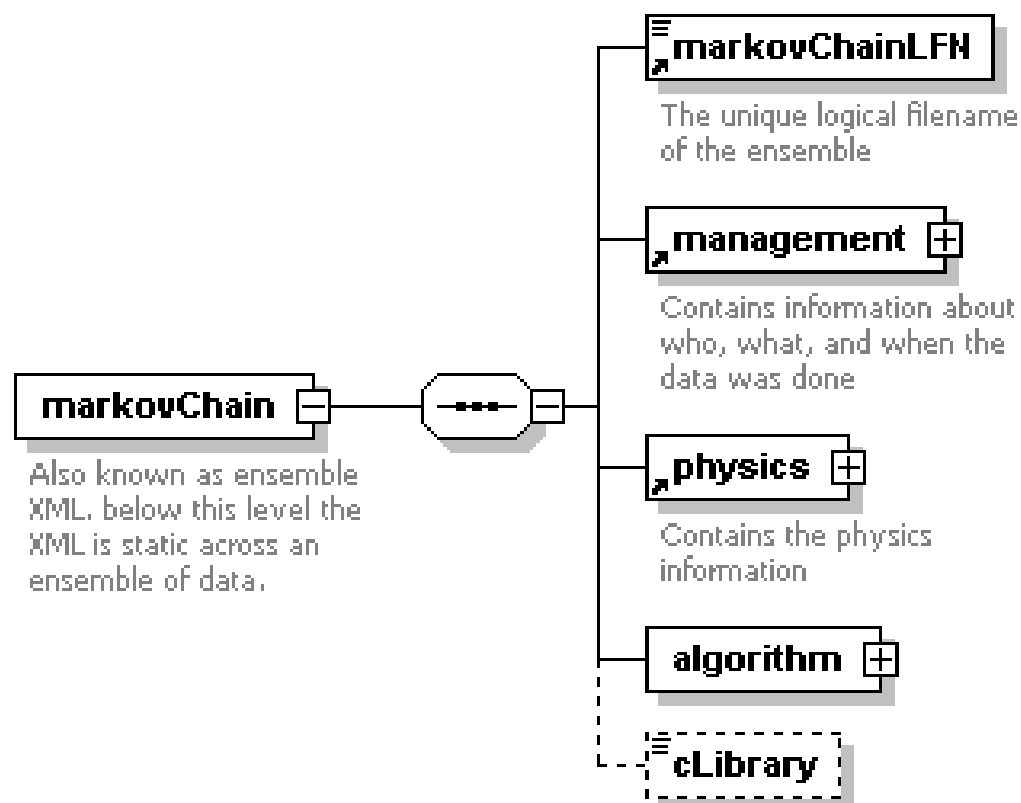
**MDWG agreed on schema: QCDml 1.0**

## Changes since Lat'03:

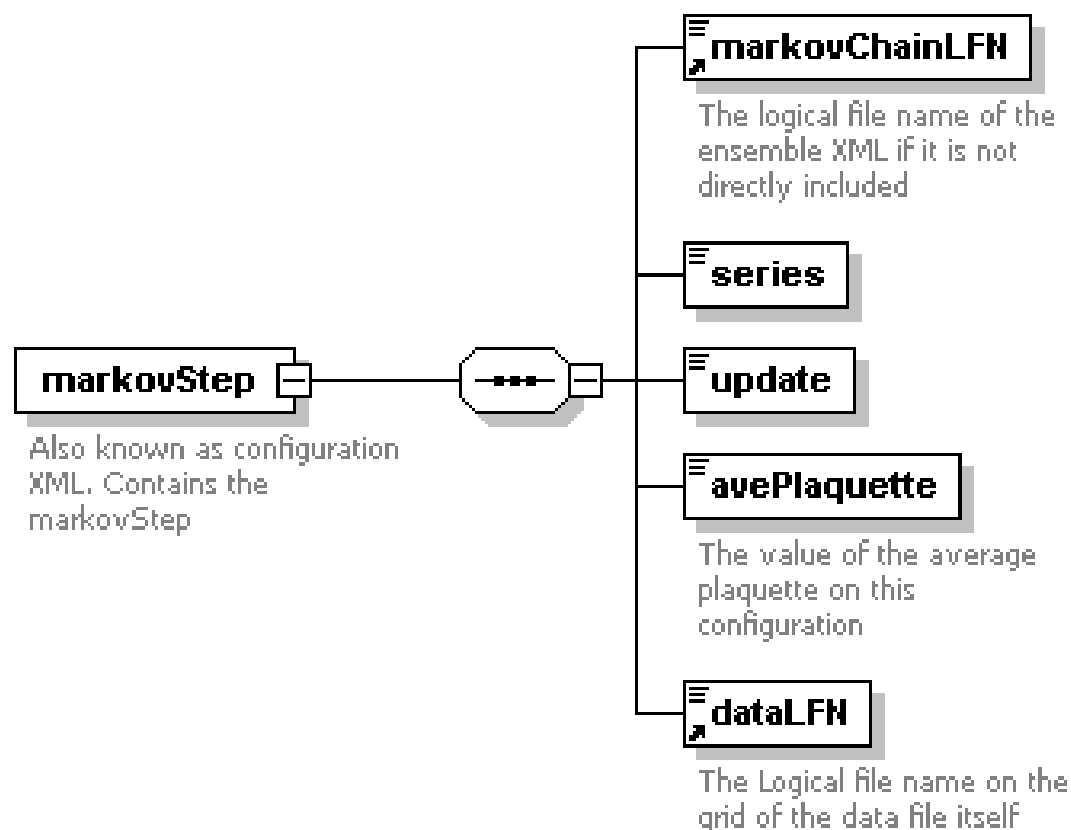
- ❑ Adopted new strategy for marking-up action
  - ❑ Split metadata into two parts:
 

Markov chain	↔	Ensemble XML
Markov step	↔	Configuration XML
- Contents remained unchanged, but usability improved

# Ensemble XML



# Configuration XML

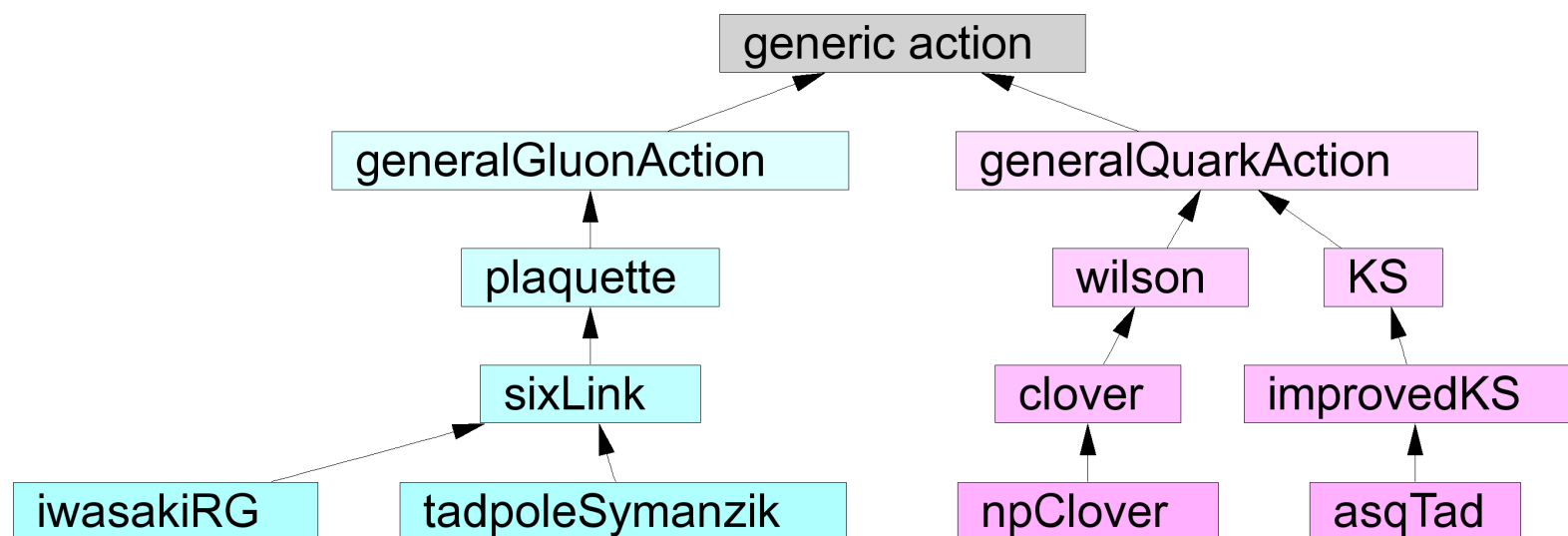


# Physics

Need to markup **action parameters** and **lattice size**.

Description of action parameters most critical to ensure metadata being **unique** and **extensible**.

**Strategy: hierarchy of actions**



## Action

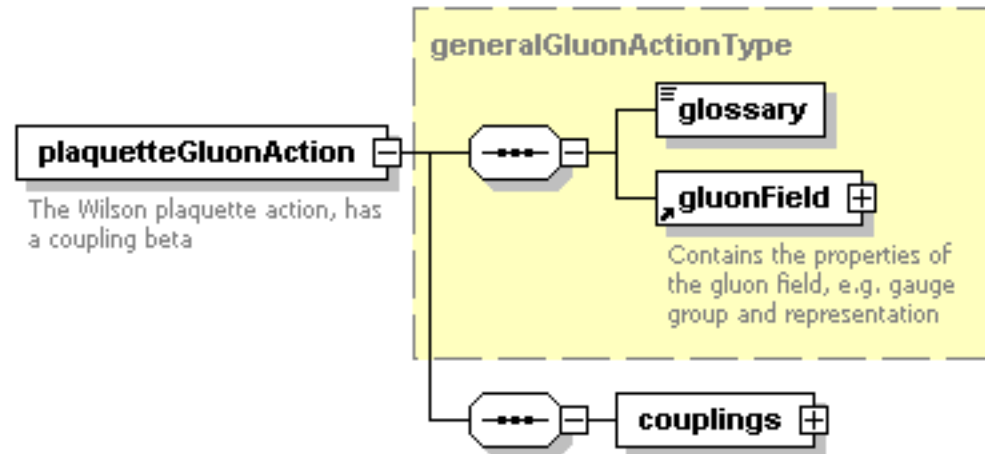
Needed for describing action:

- ❑ **Array of couplings**
    - Name and value
    - Information on  $N_f$
  - ❑ **Description of fields**
    - Normalisation
    - Boundary conditions
  - ❑ **Further information**
    - E.g. choice of couplings
- **Glossary**

glossary = URL of a (human readable) documentation provided by contributors

*Discussion: Guidelines for glossaries needed*

## Example: Wilson-Plaquette Action



```

<plaquetteGluonAction>
  ...
  <couplings>
    <elem>
      <beta>5.2</beta>
    </elem>
  </couplings>
</plaquetteGluonAction>

```

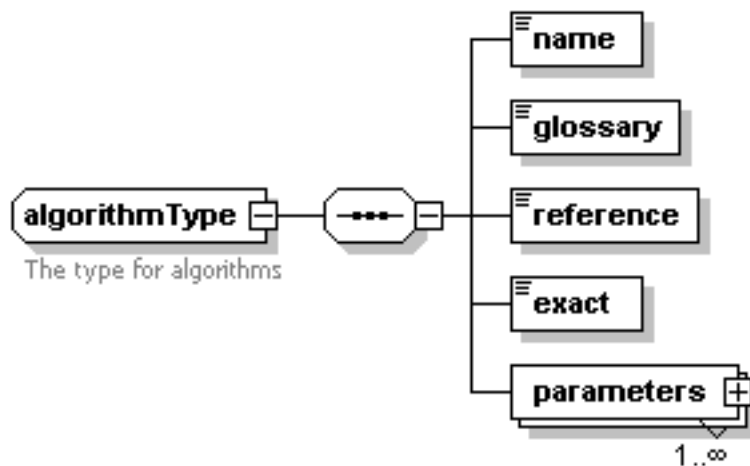
# Algorithms

Putting algorithms into a general structure is difficult

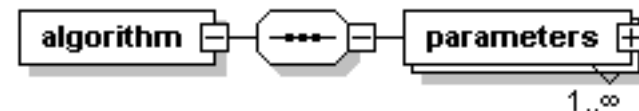
→ algorithm parameters left unconstrained

→ parameters not searchable

Ensemble:



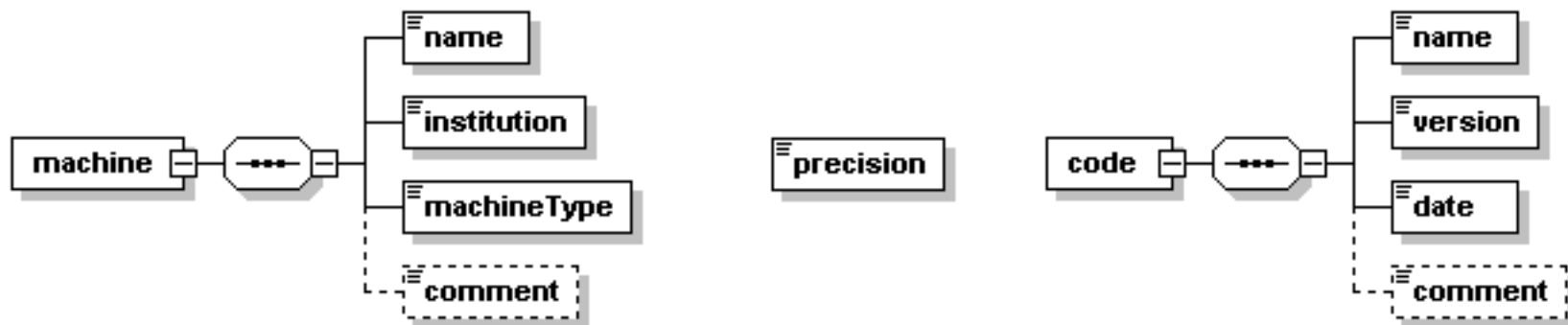
Configuration:



*Discussion: Action parameters allowed to change within ensemble*



## Machine and Simulation Code Details



- Details are part of configuration XML
- Most elements be unconstrained
- Precision  $\in \{ \text{single, double, mixed} \}$

# Data Management

## ☐ Information about contributors

collaboration name, project name

## ☐ Information needed for checking data integrity

CRC, plaque value

## ☐ Archive history

- Actions  $\in \{\text{add, replace, remove}\}$  ensemble/configurations
- Which participant acted when and why?

## Status XML Schema

Working group agreed on QCDML 1.0:

→ Check ILDG website <http://www.lqcd.org/ildg>

→ Visit QCDml tutorial next **Friday, 18:30-19:30**

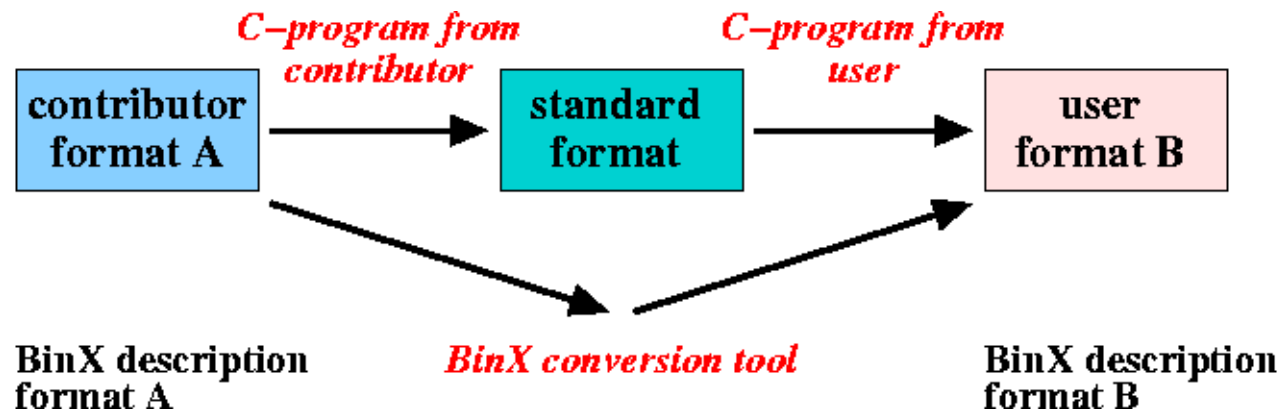
Check ILDG website for presentation

**Request for comments until end July 2004**

Email: [qcdml@rccp.tsukuba.ac.jp](mailto:qcdml@rccp.tsukuba.ac.jp)

# Binary File Format

Original strategy:



## Discussion:

- Status of standard binary format?
  - intermediate format, virtual format, mandatory format
  - avoid doubling storage requirements
- Conversion by C-library and/or BinX?
- How to package configurations (→ MWWG)?

## Open Issues

- ☐ Support gauge fixed configurations
- ☐ Extend QCDml to propagators/correlators

## Roadmap

- ☐ Final release of QCDml 1.x (end July)
- ☐ Agree on file format together with MWWG (end July)
- ☐ Complete list of actions
- ☐ Complete QCDml write-up
- ☐ Start using QCDml for production runs